Mitri Kitti, Matti Pihlava, and Hannu Salonen

Search in Networks: The Case of Board Interlocks

Aboa Centre for Economics

Discussion paper No. 116

Turku 2017

The Aboa Centre for Economics is a joint initiative of the economics departments of the University of Turku and Åbo Akademi University.

ISSN 1796-3133

Printed in Uniprint
Turku 2017

Mitri Kitti, Matti Pihlava, and Hannu Salonen

Search in Networks: The Case of Board Interlocks

Aboa Centre for Economics
Discussion paper No. 116
June 2017 (first draft November 2016)

ABSTRACT

We present a model for the dynamics of networks in which edges represent positions in organizations, holders of which are connected to each other when the positions belong to the same organization. Once a vacancy is opened, the new employee can be hired from the current network. In particular, it is possible that the further away a candidate is in the network from the firm having the vacant position, the less likely it is that the candidate is chosen. The search may also involve preferential attachment in the sense that people with high numbers of positions are more likely to be chosen. Microeconomic foundations of the search process are presented. An empirical application to a board interlock network demonstrates that the model is capable of explaining how such networks are formed. It is observed that distances from firms to candidates and the candidates' numbers of positions drive the process.

JEL Classification: D83; D85; Z13

JEL Classification: Keywords: search, two-mode network, network formation, board interlock network

Contact information

Mitri Kitti (corresponding author)
Department of Economics
University of Turku
FI-20014, Finland
Email: mitri.kitti (at) gmail.com

Matti Pihlava
Email: majupi (at) utu.fi

Hannu Salonen
Email: hannu.salonen (at) utu.fi

Acknowledgements

Funding from the Academy of Finland is gratefully acknowledged. We thank seminar participants at the University of Konstanz, Copenhagen Business School, and the participants of the XXXVIII Annual Meeting of the Finnish Economic Association and XXXVI Sunbelt Social Networks Conference for their comments.

1. Introduction

In general, it is widely accepted that social networks play an important role in labour markets.* However, quantifying their impact and assessing how exactly do the connections and peoples positions in the network affect their labour market outcomes is still a major challenge. The question addressed in this work is how the network channel in recruiting new people functions. The main obstacle for studying this question is that the topology of social networks is rarely observed at the scale that would allow for fitting and testing any explicit model for role of networks. Boards of companies as well as board appointments, on the other hand, are well documented. Hence, although the labour market of corporate elite involves less than a percentage share of total labour force, it is ideal for studying the network channel in the search of new employees.

One of the challenges in analyzing the role of social networks in labour markets is the endogeneity arising from the present network affecting the future connections between agents. For example, a candidate who is found in the social network through a small number of intermediary persons, e.g., through referrals, may get hired, which in turn creates new ties between people and affects their prospects of getting new jobs in the future. Hence, the formation of networks is the key to understanding the properties of observed networks.

In this work the attention is on modeling the dynamics of two-mode networks in which there is a fixed number of positions, holders of which are connected to each other when the positions belong to the same organization. In principle, there are two sets of nodes representing different social entities: the individuals and the organizations. For example, in a board interlock network, which is our main application, one set of nodes is composed of board members, and the other of firms. These kinds of networks are known as two-mode networks or affiliation networks. The networks analyzed in this work are assumed to have more structure than just the division of agents into groups and individuals. The additional element is the set of positions or posts that is assumed to remain the same in time. In essence, each group has a given number of positions that can be occupied by the individuals of the network.

There is a variety of models for network formation and evolution, for an overview see Jackson (2005) and Newman (2003). Our approach relies on a search and matching type process, where once a vacancy is opened, the new

*For an early work on the importance of social networks in the job search process see Granovetter (1974).

holder of the position is searched from the current network or the vacancy is filled by choosing a person outside of the holders of current positions in the network. Hence, the search process presented in this work resembles the models of Jackson and Rogers (2007) and Bramoullé et al. (2012) that involve search of new connections in the neighbourhoods of randomly met nodes. In particular, the search is affected by the distance of people in the network from the firm having the vacancy and the number of positions that people have. The distance effect can result from a referral type of exchange of information. The number of positions, on the other hand, may signal the productivity of a candidate.

The search and matching in labour market networks has been widely studied in economics literature, the topic is surveyd by Ionnides and Loury (2004) and Topa (2011). For classical contributions on wage determination in search markets see Diamond (1981), Montgomery (1991, 1992), and Mortensen and Vishwanath (1994). There is also substantial literature on referrals as a method for finding workers, see Dustmann et al. (2016) and Galenianos (2013) for recent works.

The search process presented in this paper has a number of distinguishing features that make it different from network search models presented in the previous literature. First, the focus is on the evolution of the network, while in many of the labour market search models the actual network structure plays no role or it is taken as exogenously given (Calvó-Armengol and Zenou, 2005; Fontaine, 2008), see, however, Boorman (1975), Calvó-Armengol (2004), and Galeotti and Merlino (2014) for notable exceptions. Second, a crucial element of a two-mode network is that an individual may hold several positions in different organizations at the same time. Page and Wooders (2007, 2010) have studied the formation of this kind of networks in which the other mode represents clubs and the individuals choose strategically their club memberships. The third feature in our approach is that the set of individuals is not fixed; only the firms and the positions are assumed to stay the same.

Our model is as such applicable to any two-mode network where the available positions are fixed. This situation is typical for representative bodies of various organizations such as boards of companies or associations, editorial posts of journals, or various bodies of parliamentary systems (Porter et al., 2005). Collaboration networks, in which each member of a team has a particular role (Uzzi and Spiro, 2005), can be treated as two-mode networks with fixed numbers of available positions. In this work the attention is on board interlock networks, which have raised a lot of attention in social sciences beginning from the works of Galaskiewicz and Wasserman (1981) and Mintz and Schwartz (1981).

From an economics point of view board interlock networks or interlocking directorates are interesting for several reasons. Social networks affect market outcomes (Granovetter, 1973; Rees, 1966)—in particular, as argued by Saloner (1985), networks improve the quality of the director-management match, and hence corporate performance. On the other hand, social networks may have a detrimental effect to corporate governance (Hallock, 1997; Fich and Shivdasani, 2006; Kramarz and Thesmar, 2013), while at the same time there is evidence that companies with highly networked boards may perform better (Larcker et al., 2013). The role of networks on the careers of people in top corporate positions has been documented in several studies; e.g., Engelberg et al. (2013) and Liu (2014) show a statistical relation between labor market prospects (earnings, turnover) and connectedness, Berardi and Seabright (2011) present a dynamic model for the coevolution of careers and professional networks, and Lalanne and Seabright (2016) find a causal impact on connections to remuneration.

The main advantage of a structural model, such as the one presented in this paper, is that it allows for carrying out counterfactual analysis. For instance, using our model it is possible to compare the outcomes when the distance of a candidate to a firm with a vacant position has either no effect at all or the search is restricted to firms' immediate neighborhoods. This is related to the question whether the role of board interlock networks is to transmit information regarding the quality of potential board members or whether the connections in the network just allow for an exchange of favors. In the latter case, one might expect that favors are exchanged only between board members that directly know each other, implying that the new board members hired through the network channel would mainly come from firms' immediate neighborhoods. Our empirical findings support the former explanation that the network transmits information on candidates quality; there is no exceptional peak at board nominations from distance one, and a model where the variance of candidates' productivity signal is linear as a function of distance leads to a good fit for the observed distance distribution.

This paper is structured as follows. Section 2 introduces the two-mode networks with given set of positions and the basic concepts used in the paper. The model for the evolution of the network is presented in Section 3, and its microeconomic foundations are studied in Section 4. We also discuss the extensions of the model to a more standard labor market setting where some people are unemployed and the search is carried out through connections of people in a standard one-mode network. Fitting the model to data is discussed in Section 5. Section 6 presents an empirical application which demonstrates that the model is able to capture the main features of a real world network and matches the observations that new board members are found relatively close to the current board in the interlock network and they tend to hold multiple positions. Conclusions are discussed in Section 7.

2. Two-Mode Networks with Firms and Individuals

In a two-mode network there are two kinds of actors. In this paper these actors are assumed to be firms and individuals. There is a finite set of firms F with given positions V. Each position $v \in V$ is assumed to be held by some individual. The set of individuals is denoted by I. Each position belongs to some firm, let f_v stand for the firm having the position $v \in V$. The positions belonging to the firm $f \in F$ are $VF(f)$, and the positions held by the individual $i \in I$ are $VI(i)$. Individuals holding positions in firm f are denoted by the set $I(f)$.

Two positions are linked to each other if the same person holds them. Such persons are called interlockers. In particular, each $v \in V$ corresponds to a list of other positions held by the holder of v. Let V_v denote this list. In the following $n_v = |V_v|$ is the total number of positions held by the person who occupies the position v. It is assumed that if a position in a firm is held by some individual, all the other positions of that firm are occupied by some other people.

As an example, consider a network of nine positions labelled with $v^1, \ldots, v^9$ and the set of firms $F = \{a, b, c\}$. The positions $v^1, \ldots, v^3$ belong to firm a, v^4 and v^5 belong to firm b, and the rest of belong to c. Hence, $f_{v^j} = a$ for $j = 1, \ldots, 3$, $f_{v^j} = b$ for $j = 4, \ldots, 5$, and $f_{v^j} = c$ for $j = 6, \ldots, 9$. Moreover, assume that there are seven individuals $I = \{1, \ldots, 7\}$ such that the first three have a position in firm a, the last four have a position in firm c, and individuals 1 and 4 hold also positions in firm b. The two-mode network corresponding to this example is illustrated in Figure 1. Note in particular that individuals 1 and 4 are interlockers with $VI(1) = \{v_1, v_4\}$ and $VI(4) = \{v_5, v_6\}$ or $V_{v^1} = V_{v^4} = \{v^1, v^4\}$ and $V_{v^5} = V_{v^6} = \{v^5, v^6\}$.

Note that in essence each position corresponds to an edge between a firm and an individual. What is crucial is that the number of these edges is limited, because each firm as a limited number of positions. When it comes to the particular application on board interlock networks, the board size or its upper and lower bounds are usually set in the corporate by-laws. Hence, the number of positions in board interlock networks is relatively stable, but different companies may have different number of board members.†

†The size of the board is known to be related to the firm performance, see, e.g., Cheng (2008).

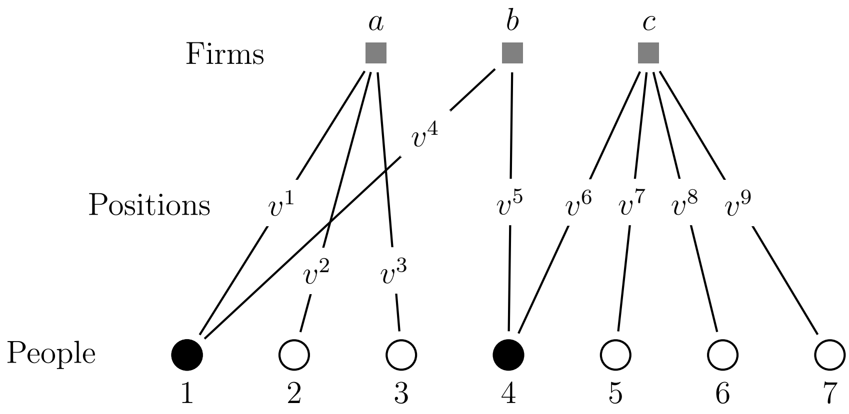


Figure 1: Example of a two-mode network with positions as edges. The interlockers are indicated with black circles.

The holders of positions form a network; people having positions in the same firm are connected to each other, and two firms sharing a common individual are connected to each other. To describe the network composed of positions and their holders let us first consider the neighbourhoods of available positions in the network.

The immediate neighbours of $v \in V$ consist of all positions in the firm f_v, i.e., it is the set $N_0(v) = VF(f_v)$. The one-step neighbourhood of v contains all the positions in the firms which have board interlocks with the firm;

$$N_1(v) = \cup\{VI(i) : i \in I(f_v)\}.$$

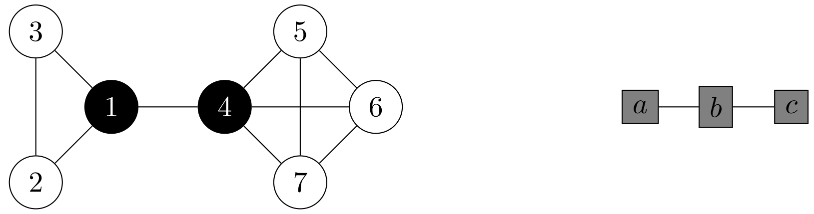
Higher order neighbourhoods, the k-step neighbourhoods for $k \geq 2$, are obtained the same way;

$$N_k(v) = \cup \{VI(i) : i \in I(f_w),\ w \in N_{k-1}(v)\}.$$

In other words, $w \in N_k(v)$ corresponds to at most k interlockers through which v is connected to w.

To clarify the concept of a k-step neighbourhood let us continue the previous example. Regardless of the holders of the positions, the zero-step neighbourhoods are $N_0(v^j) = \{v^1, v^2, v^3\}$ for $j = 1, 2, 3$, $N_0(v^j) = \{v^4, v^5\}$ for $j = 4, 5$, and $N_0(v^j) = \{v^6, \ldots, v^9\}$ for $j = 6, \ldots, 9$. The k-steps neighbourhoods are

$$\begin{aligned}
&N_1(v^j) = \{v^1, \ldots, v^5\} \text{ for } j = 1, 2, 3,\\
&N_1(v^j) = V \text{ for } j = 4, 5,\\
&N_1(v^j) = \{v^4, \ldots, v^9\} \text{ for } j = 6, \ldots, 9,
\end{aligned}$$



a) Projection to individuals   b) Projection to firms

Figure 2: Example of projections of a two-mode network.

and $N_2(v^j) = V$ for all $j = 1, \ldots, 9$.

Let $N(v)$ denote the list of all non-empty neighbourhoods of $v \in V$. To be specific, $N(v) = (N_0(v), N_1(v), \ldots, N_{n-1}(v))$. A neighbourhood configuration at time t is the ordered list $\{N(v)\}_{v \in V}$. Because there are finitely many positions, there are also finitely many neighbourhood configurations. The set of all possible neighbourhood configurations is denoted by $\mathcal{N}$. The projection of $N \in \mathcal{N}$ to the neighbourhoods of v is $N(v)$ with the corresponding space $\mathcal{N}_v$.

If w is found in some k-step neighbourhood of v but is not found in any j-step neighborhood for $j < k$, then the distance between v and w is $d_{w,v} = k$. If v is not found from any k-step neighbourhood of w, then $d_{w,v}$ is set to d^∞. In the previous example $d_{v^1,v^j} = 1$ for $j = 4, 5$, and $d_{v^1,v^j} = 2$ for $j \geq 6$.

A two-mode network can be transformed or projected into networks corresponding to the different modes. For example, the network in Figure 1 can be projected into a network of individuals and a network of firms. These projections are illustrated in Figure 2. Observe in particular that the projection into individuals is basically a social network of people affiliated with the firms. In principle, we could operate with such one-mode social networks when considering the search of a new person to a vacant position. This would, however, be somewhat more complicated in the context of our main application. Namely, in board networks it is essential how many positions a person has, and how far from a company does the new person come from. For a social network, such as that illustrated in the left panel of Figure 2, these relations are not apparent anymore.

The main reason for working with a two-mode network rather with its one-mode projection into a network of individuals is that the opening and filling of vacant positions will simply mean rewiring an edge in the two-mode network and it is easy to keep track of peoples' affiliations. For example, assume that person 6 in Figure 1 is replaced. In the projected network of the left panel of Figure 2 this would mean that node 6 is removed and all the edges to that node are to be rewired to the same node. Hence, instead of rewiring only one edge we would have to take care of rewiring all the edges there were to a removed node, and we would also need to keep track of peoples affiliations to different companies — information that is apparent in the two-mode representation.

We emphasize that although it is more convenient to describe our model for two-mode networks, it can be extended to ordinary one-mode social networks by keeping track of peoples' affiliations and taking care that all relevant edges are rewired when vacancies are filled. Note in particular that the distance between two positions could be expressed in terms of the distance between individuals in the projected network. For example, the distance between v^2 and v^4 in Figure 1 can be obtained by taking the distance between individuals 2 and 5 in the projected network of individuals and subtracting one from this number, .i.e., the distance is $4 - 1 = 3$ as in the original two-mode network.

3. Evolution of Two-Mode Networks through the Filling of Vacancies

This section presents a model for the dynamics of two-mode networks with a given set of positions that are opened and filled as time goes on. As described in the previous section, the network is essentially determined by its neighbourhood configuration at each time instant. The purpose is to present simple probabilistic rules updating the neighbourhood configuration. It is assumed that time is discrete and it is indexed with $t = 0, 1, 2, \ldots$. All variables that depend on the network, such as neighbourhoods of positions, are indexed with t, whenever needed.

The model for the evolution of two-mode networks is essentially a stochastic process for the edges and vertices of the network. The edges correspond to firm-individual pairs in which the individual holds a vacancy in the firm, and the vertices correspond to firms and individuals. Most of the existing models for network formation involve only dynamics of edges, while the set of possible vertices is kept fixed. Alternatively, it is often assumed that the network is growing as in the preferential-attachment model of Barabási and Albert (1999).

In our model there is no limited pool of people holding the positions. When time goes on arbitrarily many people may have held a position. However, the network is not growing because the number of positions is kept fixed. In essence, the opening and filling of vacancies corresponds to rewiring of edges of a two-mode network. See Evans and Plato (2007) for edge rewiring with a fixed set of agents, and Lafond (2015) for a model where the number of vertices in one of the modes is fixed while in the other mode the vertices are taken from a finite population.

An important assumption in out work is that the previous period network determines the probabilities of getting chosen to a vacant position. Hence, we do not assume that there is any longer memory in the affiliations. For example, a person who has a position only in firm a in year 1, and a position only in firm b in year 2 is not creating an interlock between firms a and b. Moreover, these past connections are not assumed to affect how the network evolves in the future.

3.1. Opening and filling of vacancies

It is assumed that vacancies open at random, the probability for position $v \in V$ to open in period t is $P_o(N^t(v))$. Note that the dependency on $N^t(v)$ means that the probability of opening the vacancy may depend on the number of positions that the previous holder of the position has had or how many people the holder is connected to. If none of the positions held by the holder of v is opened then the list of positions held by this individual remains the same; $V_w^{t+1} = V_w^t$ for all $w \in V_v^t$.

In practice, opening a vacancy means that someone leaves a position. In this work we do not consider the possibility of opening completely new posts. As mentioned it is assumed that vacancies open at random, while in practice there can be different reasons for someone to leave a position; people move from one job to another, they are kicked out, or they are obliged to leave after holding the position for a certain period of time. When a firm decides to replace someone, there are typically costs—especially search costs—related to such a decision, and these costs certainly affect the probability P_o. However, in this work we do not explicitly consider the microfoundations of P_o.

When a vacancy opens, its holder is changed to a new person taken either outside of the current holders of available positions in the network or from the set of holders of other positions. The first case leads to $V_v^{t+1} = v$, i.e., the only position held by the new holder is v. Hence, the list of positions held by the holder of position $w \in V_v^t$ is updated such that v is removed from V_w^t. In the second case, when the new person for the vacancy v is chosen among the holders of other positions, there is some position $w \notin V_v^t$ holder of which gets the position v. In this case V_w^{t+1} is appended with v.

Recall that each position corresponds to an edge of the two-mode network. Hence, opening and filling the vacancies can be interpreted as an edge rewiring process, where the rewiring probabilities are conditional on the properties of edges, and the number and labels of nodes in the other mode of the two-mode network are not fixed.

Let us now make assumptions on the probabilities of opening and filling vacancies:

(A1) $v \in V$ opens with probability $P_o(N(v))$ such that $P_o(N(v)) > 0$ for any $N(v) \in \mathcal{N}_v$.

(A2) the vacancy is filled outside of the holders of other positions with probability $P_n > 0$,

(A3) if v is filled from the set of holders of other positions, then the probability of choosing the holder of $w \in V \setminus V_v$, is $P_h(N(w), N(v))$ such that $P_h(N(w), N(v)) > 0$ if $d_{w,v} \geq 1$, and $P_h(N(w), N(v)) = 0$ if $d_{w,v} = 0$.

The first two assumptions mean that there are two channels that can be used to fill a vacancy: the network channel described by P_o and the more conventional labor market channel described by P_n. The last assumption means that there is a search and matching process which determines the new holder of a position. The holder of w is matched with the firm f_v with a probability that depends on positions of the holder of w in the network and how these positions are related to the neighbourhoods of the vacancy to be filled. Note that having $P_h(N(w), N(v)) = 0$ if $d_{w,v} = 0$ means that v cannot be filled by any person already affiliated with the firm f_v. What is essential in the assumption (A3) is that the immediate neighbourhood of a vacancy affects how it is filled.

Let us now consider more specific functional dependencies for the probabilities P_o and P_h. First, if P_o depends only on n_v, let us denote $P_o(n_v)$, and is a decreasing function, then there is preferential attachment in the sense that a position is less likely to be opened if its holder has several other positions.

One particular functional form for P_h is obtained by assuming that P_h depends only on the distance and degree, i.e., $P_h\left(n_w, d_{w,v}\right)$. If P_h is increasing in its first argument, a person holding several positions is more likely to get the position v. Again this is a form of preferential attachment. On the other hand, if P_h is decreasing in its second argument, there is a form of peer referral; a holder of a position that is found in a close neighbourhood of the vacancy is more likely to become the new holder. In particular, when the new person is found from distance one, there is a triadic closure; the new holder of the position is known by someone in the same board with the previous holder of the position.

The first observation on the above dynamic system is that it describes a Markov chain over the neighbourhood configurations. Let $P(N^{t+1}|N^t, \ldots, N^0)$ stand for the conditional probability of a configuration N^{t+1} in period $t+1$. The Markov property means that $P(N^{t+1}|N^t, \ldots, N^0) = P(N^{t+1}|N^t)$ for any $N^{t+1}, N^t, \ldots, N^0 \in \mathcal{N}$. Moreover, the process is irreducible; it is possible to get from one neighbourhood configuration to any other configuration with a positive probability.

Proposition 1. *Under assumptions (A1)–(A3) the stochastic process of neighbourhood configurations is an irreducible Markov chain.*

The above result, proved in Appendix A, implies that the stochastic process over the neighbourhood configurations has a unique stationary distribution. Note that the number of possible neighbourhood configurations, although finite, can be extremely high, which means that in practice it is not the neighbourhood configurations that can be tracked but some statistics that depend on them. Irreducibility of the Markov chain implies that the time average of any statistics that depends on neighbourhood configurations converges with probability one. In particular, the time averages of distances and degrees of new board members converge.

3.2. The exponential search model

A particular functional form for the probability P_h by using the exponential function as in the multinomial logistic regression:

$$P_h(N(w), N(v)) \quad = \exp\left[a(d_{w,v}) + b(n_w)\right]/C, \tag{1}$$

where C is a constant such that summing the probabilities over the candidates is one;

$$C = \sum_{\substack{w \in V \\ d_{w,v} \geq 1}} \exp\left[a(d_{w,v}) + b(n_w)\right].$$

In Section 5 it is assumed that the functions in the exponential search model are

$$a(d) = -\alpha d, \tag{2}$$

$$b(n) = \beta \min\{n, K\}/K, \tag{3}$$

Notice that K indicates the largest relevant number of positions. If the number of positions is larger than K, the effect of is the same as if it was K.

In addition to assuming exponential function for P_h we could assume that P_o takes the exponential for as well:

$$P_o(N(v)) = p^0 \exp[c(n_w)],$$

where $p^0 > 0$ is a constant such that $P_o(N(v)) \leq 1$. However, in the specific case when P_o depends on the number of positions that the holder of v has, the empirical probabilities $P_o(n_v)$ can be estimated relatively easily and there is no need to assume any specific functional form for P_o.

The model resulting from equations (2)–(3) is flexible enough for producing both realistic degree distribution and networks with small-world properties such as high clustering coefficient. In particular, a large value of β would imply a high probability of vacancies filled with people having a large number of positions, which is likely to lead to a small-world network. On the other hand, a large value of α would imply a high probability for the new person to be close to the holders of the positions in the same firm, which is likely to lead to a highly clustered network.

The exponential form in equation (1) resembles other probabilistic models for network formation such as the stochastic actor-oriented model of Snijders (2001) and exponential random graphs (Wasserman and Pattison, 1996). However, we emphasize that the exponential search model is not a specific case of either of these two families although they are applicable for two-mode networks (Koskinen and Edling, 2012; Wang et al., 2009). It is also worth observing that similarly as for stochastic actor oriented models or exponential random graphs, it is possible include node specific affects into the model leading to a more general formulation:

$$P_h(w,v) \quad = \exp\left[\sum_i \theta_i g_i(w,v)\right]/C,$$

where function g_i describe the characteristics of holders of positions w and v. For example, characteristics of w and v that could be of interest are the industries of the firms having the positions, or the age and gender of the holders of w and v. On the other hand, we could also include properties related to the pair (v,w), such as an indicator whether the person holding w has previously been in a board that has nominated any person in the board of firm f_v. This kind of favor exchange can be one of the motives to nominate certain people.

4. Microeconomic Foundations

In this section we present simple microeconomic foundations for the probabilistic model that describes the filling of vacancies. It is assumed that when a firm opens a vacancy and searches for a person to fill it, there are two options: either the firm hires someone who already has a position or someone who is not holding any other position. In the former case the new person can either be found from searching through then neighbourhoods of the holders of the positions of the firm, or outside of any of the neighbourhoods.

The microeconomic foundations are build upon productivity signalling, where the quality of the signal depends on the distance of two positions and on the number of positions that people have. We emphasize that there may also be other forces that affect the evolution of the network such as favor exchange. However, in this work we only focus on the two simple features related to the topology of the network: the distance and the number of positions.

4.1. Productivity signalling

As is typical in the search and matching literature, it is assumed that the productivity of each person is unknown but the firm gets a private signal on it. Note that in the framework of this paper, productivity can also be interpreted simply as a variable describing how suitable a person is for carrying out the task related to a vacancy. Let $s(n, d) \geq 0$ stand for the signal of the productivity of person holding $n \geq 1$ positions and located on the distance $d \geq 1$ from the firm. Recall that n_v stands for the number of positions held by the holder of position v.

It is assumed that the payoff from hiring a person with productivity $z \geq 0$ is $u(z)$ for each firm. To be specific, u is the firm's von Neumann-Morgenstern utility function. In economics it is common to assume that firms are risk neutral since the shareholders can efficiently diversify away any firm-specific risks via the stock market. In that case u would be a linear function of z. In particular, when it comes to firms' investment decisions risk neutrality is a prevailing assumption. However, there are several reasons why firms may behave in a risk-averse manner. The ownership can be concentrated, which is especially the case for family businesses (Czarnitzki and Kraft, 2009), or the control is delegated to risk averse managers (Brenner, 2015). We also emphasize that the choice of a new board member is not necessarily comparable to an investment decision. Hence, even a publicly listed company may exhibit risk aversion when it comes to the choice of a new board member.

When a firm receives a signal $s(n, d)$, the productivity z becomes a random variable with expected value $s(n, d)$ and variance determined by n and d. The signals are drawn from the productivity distribution that may depend on n and d. However, the ex-ante distribution, i.e., distribution of a signal drawn for a randomly selected person, is the same as the productivity distribution of the whole population of all holders of positions. In the extreme case when the signal is completely uninformative, the productivity distribution is the same as the distribution of productivity after the signal.

The firms' expected payoff is $\mathbb{E}[u(z)|s(n_w, d_{w,v})]$ from hiring the holder of position $w \in V$ to the vacancy v. The firm is assumed to know the distribution of signals and productivity. In practice, the firm may obtain the information through communicating with the peers of the person affiliated with the firm. As a result, a candidate may get a referral for the job.

In the simplest case, the search for a person to a vacancy is costless, which entails that the firm hires a person with the highest expected payoff. Let s^1 denote the highest signal of a person outside of holders of current positions. It is assumed that s^1 is drawn from some distribution known by the firm.

Let us briefly consider the other side of the market. It can either be assumed that people who are offered a position always accept the offer, or they can reject the offer and wait for new offers to arrive. When all the firms are similar and there is no cost of holding multiple positions at the same time, it can be assumed that a person always accept the offer. For simplicity let us assume that this is the case. Note that in the latter case, a person may find it better to reject an offer and wait for a new offer if it is likely that some "better" firm will make one in the future. In other words, when it is costly to hold many positions and firms are different, for instance larger firms pay higher salaries, there can be an option value in waiting for new offers.

Let us now return to the filling of vacancies under the assumption that offers are always accepted. In this case the firm hires a new person to v with signal s^1 outside of holders of positions in the network, when

$$\mathbb{E}[u(z)|s^1] > \max\{\mathbb{E}[u(z)|s(n_w, d_{w,v})] : w \in V\}.$$

On the other hand, the holder of position $v' \in V$ is hired if the choice v' maximizes $\mathbb{E}[u(z)|s(n_w, d_{w,v})]$ over $w \in V$ and

$$\max\left\{\mathbb{E}\left[u(z)|s(n_w, d_{w,v})\right] : w \in V\right\} \geq \mathbb{E}\left[u(z)|s^1\right].$$

If there are multiple maximizers we can assume that the new person is taken randomly among them.

The search and matching process as described above determines the probabilities P_n and P_h. To be specific, P_n and P_h are the ex ante probabilities of hiring a person outside of the network or a person who already holds one of the positions in the network, i.e., probabilities prior to a firm receiving any signal. The distributions of signals affect the properties P_h. To analyze these properties some concepts are needed. The first is a mean preserving spread. The signal s^1 is said to be mean preserving spread of s^2 if the distribution of s^1 is the same as the distribution of $s^1 + q$ where $\mathbb{E}[q|s^1] = 0$. This means that the expected value of s^1 is the same as the expected value of s^2 but the variance of s^1 is at least the same as the variance of s^2. When s^1 is a mean preserving spread of s^2, we can say that s^2 is a more informative signal.

The second concept that is needed is the first order stochastic dominance. The signal s^2 has the first order stochastic dominance over s^1 if there is a random variable q such that $q \leq 0$ and the inequality is strict at least in one state, and the distribution of s^1 equals the distribution of $s^2 + q$. In other words, the expected value of s^1 is lower than the expected value of s^2 while the variance of s^1 is at least the same as the variance of s^2.

Assume that the information on candidates comes through referrals. In that case the information on candidates becomes less reliable the more there are intermediaries in the communication, which results on the signal becoming less informative when the distance increases. This property is captured in the mean preserving spread; the signal s^1 is less informative than s^2 if s^1 is a mean preserving spread of s^2.

If the firm is risk averse and the signal becomes less informative the further away the vacancy is from the firm, then the resulting probability of a match is a decreasing function of the distance. The proof is presented in Appendix A.

Proposition 2. *Assume that* $n_{v^1} = n_{v^2}$ *for* $v, v^1, v^2 \in V$, $v \neq v^i$, $i = 1, 2$, *and* $v^1 \neq v^2$. *If the firm is risk averse and the distribution of* $s\left(n_{v^1}, d_{v^1,v}\right)$ *is a mean preserving spread of the distribution of* $s\left(n_{v^2}, d_{v^2,v}\right)$ *when* $d_{v,v^1} > d_{v,v^2}$, *then the probability* P_h *of the holder of* v^1 *being chosen to the vacancy* v *is no larger than the corresponding probability for the holder of* v^2.

The number of positions held by a person may signal the productivity of a candidate. In particular, having more positions may signal a person's ability and hence impact positively the productivity signal. One explanation to higher productivity due to being in several boards comes from the access to more information (Mizruchi, 1996). On the other hand, there is also a mechanism acting to other direction; a high number of positions may signal low productivity of the candidate when assigned to one more board. In either case, the first order stochastic dominance can be used in capturing the effect of number of positions ("degree") in the productivity signal. The proof of the following result can be found from Appendix A.

Proposition 3. *Assume that* $d_{v^1,v} = d_{v^2,v}$ *for* $v, v^1, v^2 \in V$, $v \neq v^i$, $i = 1, 2$, *and* $v^1 \neq v^2$. *If the firm is risk averse and the distribution of* $s\left(n_{v^2}, d_{v^2,v}\right)$ *has the first order stochastic dominance over the distribution of* $s\left(n_{v^1}, d_{v^1,v}\right)$ *when* $n_{v^2} > n_{v^1}$ *(alternatively when* $n_{v^2} < n_{v^1}$*), then the probability* P_h *of the holder of* v^1 *being chosen to the vacancy* v *is no larger than the corresponding probability for the holder of* v^2.

4.2. Foundations of the exponential search model

In this section we consider the exponential search model of Section 3.2. The purpose is to give simple assumptions for firms' preferences and private information that lead to the exponential search model. As will be shown, the probability of the form (1) for filling a vacancy arises when the firms are assumed to have a common constant absolute risk aversion utility (CARA) utility function, the productivity is normally distributed, and the firms get private signals on productivities of each candidate that are type I extreme value distributed.

As before, I is the set of candidates available for a vacancy. The set of firms is F. In the following n_i stands for the number of positions held by $i \in I$, and $d_{i,k}$ stands for the distance of $i \in I$ from the firm $k \in F$. Note that $n_i = |V_{v^i}|$ where v^i is any position held by i, and $d_{i,k} = d_{v^i,v^k}$, where v^k is a vacancy that is to be filled in the firm $k \in F$.

Assume now that the expected value μ of productivity z depends on the number of positions, and the variance σ^2 depends on the distance such that

$$\mu = A + b(n_i) + \varepsilon_{i,k}, \tag{4}$$
$$\sigma^2 = -2a(d_{i,k}), \tag{5}$$

where $A \in \mathbb{R}$ is a constant and $a(d_{i,k}) \leq 0$. The term $\varepsilon_{i,k}$ reflects the private information that firm $k \in F$ has on the productivity of individual $i \in I$. These terms are not observed by an outsider but they affect the choices made by the firms. Note also that when b is an increasing function, then the expected productivity is increasing in the number of positions held by individuals. Moreover, when $a \leq 0$ is decreasing, the higher the distance the larger the variance of the productivity.

It should be observed that this model can be interpreted in terms of productivity signalling. The expected value μ is the expected value of the productivity conditional on the number of positions, and σ^2 is its variance. In principle, the firm receives a signal on the candidates productivity as before, but now the signal has a private component $\varepsilon_{i,k}$. The main assumption is that the signalling leads to a conditional productivity distribution that depends on $\varepsilon_{i,k}$, $d_{i,k}$, and n_i with mean and variance as in (4) and (5).

The following proposition gives conditions under which the probability of choosing a candidate with a certain number of positions and distance to the firm is determined by the logistic formula of equation (1). The proof of the result is presented in Appendix A.

Proposition 4. *Assume that*

1. *the utility function of the firm $k \in F$ is the CARA utility $1 - exp(-\gamma z)$ for $\gamma > 0$,*
2. *the productivity z is normally distributed with mean and variance as in (4) and (5), and*

3. $\varepsilon_{i,k}$, $i \in I$, *are independent identically distributed random variables from type I extreme value distribution that are known by the firm.*

It follows that the conditional probability of filling a vacancy P_h *is of the form*

$$P_h\left(n_i, d_{i,k}\right) = \exp\left[\gamma a(d_{i,k}) + b(n_i)\right]/C, \tag{6}$$

where $C > 0$ *is the normalizing constant.*

The above result is important, because it means that there is a simple structural model underlying the exponential search model. Note that the same result as above is obtained alternatively by assuming a CARA utility $1 - exp(-z)$ and a variance $-2\gamma a(d_{i,k})$. Hence, the variance and the constant of risk aversion γ cannot be disentangled in this structural model.

Proposition 4 offers an interpretation for the parameters α and β of the model corresponding to equations (2) and (3). The parameter β is the rate at which the productivity decreases or increases as a function of the degree $|x|_K/K$. The parameter α, on the other hand, describes the rate at which the "risk" associated with the productivity signal changes when the distance of a candidate from the firm increases. Note that α in (3) could be interpreted directly as the coefficient of risk aversion γ when the variance is equal to the distance; $\sigma^2 = d$. Alternatively, α could be interpreted as the multiplier in the variance term $\sigma^2 = \alpha d$ while $\gamma = 1$. Hence, the estimated α is basically a composition of the coefficient γ and the rate of increase in the variance when the distance increases.

4.3. Extensions and modifications

In the model discussed in previous sections, it is assumed that all the offers are accepted and all the positions are filled. However, in general this is rarely the case in labor markets. In this section we briefly discuss how the model can be extended to the more standard case where there are vacancies and unemployed people at the same time.

Recall that there are two channels in recruiting new people in our model: the network channel, which we mainly focus on, and the more conventional channel. Which of the two channels is used is determined by the probability P_n. When the new people comes outside the network we could use the usual search and matching technology; the expected number of positions filled $M_1(U, Y)$ depends on the number of people U searching for a job and the number Y of vacancies that are filled outside of the network.

To allow for unemployed job seekers in the network, it should be assumed that the network N no longer represents a two-mode network but a usual one-mode social network of ties between people. As outlined in the end of

Section 2 this modification is straightforward, but requires changing the way how edges are rewired. The second matter is that we need to keep track of peoples affiliations and the employment status. For simplicity, let us assume that the network N carries this information, and let V stand for the vacancies that are filled through the network channel. Moreover, now the distance is between individuals rather than firms as outlined in the end of Section 2, and the degree is the number of ties that a person has.

In the previous sections it is assumed that all companies know the pool of people in the network N and get signals on the productivity of each of them. However, we could assume that each firm has access only to the component of the network that includes the workers of the firm. Note that the network may be fragmented into several components such that different firms can reach different people through the network channel.

Recall that each firm searches for people that maximize the expected payoff $\mathbb{E}[u(z)|s(n,d)]$. There may, however, be a reservation payoff level for hiring anyone, and this payoff level can be determined by the option value of waiting for the next period and searching again, i.e., it may depend on N, Y, and V. Given that the reservation payoff is high enough, there will be a positive probability $P_e(N,Y,V)$ that a vacancy is not filled, and this probability determines the expected number of people who are found through the network channel $M_2(N,U,Y,V)$.

By combining the two channels of finding new people, it is possible to derive a network dependent matching function $M(N,U,Y,V) = M_1(U,Y) + M_2(N,U,Y,V)$ from the basic elements of our model. Recall that U is the number of unemployed people outside of the network N, Y is the number of vacancies that are filled outside of the network, and V is the set of vacancies filled through the network channel. Moreover, the network N carries the information on the employment status and the affiliation of each person. In particular, N relates each vacancy to a certain component of the network corresponding to the individuals that a firm can reach through the network channel. Notice that Y and V are random variable determined by P_n and P_o. In essence, P_o determines the rate of job destruction and P_n allocates the vacancies to the two recruitment channels.

Finally, it is important to observe that in our model the network does not keep memory of past affiliations or past contacts. However, in practice the social network may have memory. For example, past affiliations may affect future referrals. Accounting for the impact of past affiliations would require modeling how the ties between people decay in time. In essence, we are assuming immediate decay in this work. One possible way to incorporate time decay of ties would be to assume that the past affiliations or ties are forgotten with some probability. This extension is, however, beyond the scope of this work.

5. Fitting the Model to Data

Assume that instead of observing neighbourhood configurations there is some variable y that depends on the neighbourhood configuration and is observed at time instants $t = 1, \ldots, T$. It is assumed that $y(N^t) \in \mathbb{R}^k$. For example, y could be the "degree" distribution, in which $y_i(N^t)$ equals the number of people that are affiliated to i firms when the neighbourhood configuration is N^t.

The approach for finding the parameters that define the probabilities for filling vacancies relies on fitting the time average of variables that depend on the neighbourhood configurations into observed data. To be specific, assume that $y(N^t)$ depends on the neighbourhood configuration and an empirical realization y is observed. In practice, vector y can be computed by taking the time average of variables over periods $t = 1, \ldots, T$.

Assume that the parameters of the model belong to a set $X \in \mathbb{R}$. For a given parameter value $x \in X$, the stochastic process determined by these parameter values leads to the time average $y(x) \in \mathbb{R}^k$ over the periods $t = 1, \ldots, T$. In practice this time average can be found approximately by simulating the evolution of the network from a given initial setup several times and taking the average over the simulations. The parameters can be fitted by minimizing the least squares criterion $\|y(x) - y\|^2$, where $\| \cdot \|$ stands for the usual Euclidean norm. Depending on the relevant properties that want to be fitted, different network dependent variables or distributions can be chosen. Let us next describe in more detail the particular criterion used in the empirical application of the following section.

5.1. Fitting the exponential search model

In this section the purpose is to explain how the exponential search model as specified in equations (2) and(3) can be fitted into empirical data. The same approach can, however, be applied for other functional forms, too. The relevant parameters of the model are $P_n, K, \alpha, \beta, d^\infty$, and the probabilities $P_o(n_v)$. Because there are two separate stages—opening and filling of vacancies—we present a two-step procedure to fit the model. The first step concerns the opening of vacancies, i.e., $P_o(n_v)$, and the second how the vacancies are filled, i.e., the parameters α, β, and d^∞.

From now on the number of positions that a holder of a position $v \in V$ has, will be called the degree of v. Let $V(n)$ stand for the number of positions with degree n and $V_o(n)$ the number of positions that have been opened and have had degree n. The ratio $V_o(n)/V(n)$ is the empirical frequency at which vacancies with degree n open.

Recall that the parameter K which appears in equation (3) indicates the largest relevant degree. Hence, this parameter can be taken as the largest degree observed in the data. Let us next describe how to obtain an estimate of the probability P_n for a new board member to be taken outside of the people currently in the boards. Let n_o stand for the number of vacancies opened during the period of time spanned by the data, and let n_e denote the number of new people coming outside of the network during the time period. The estimate for P_n is simply n_e/n_o.

Finally, let us turn to the question of finding the remaining parameters α, β, and d^∞. Because α is related to the effect of distance in filling the vacancies and β is related to the degree, it can be argued that these parameters should be chosen such that the distance and degree distributions of new people fit empirically observed distributions. To be more specific, let $y^1(\alpha, \beta, d_\infty)$ stand for the distribution of distances of newly recruited people from the position that they were chosen conditional on the person being found in some of the neighbourhoods of the vacancy that was filled. Moreover, $y^2(\alpha, \beta, d_\infty)$ stands for distribution of degrees of newly recruited individuals conditional on being found among the people in the boards of the previous period. Note that y^1 and y^2 are vectors containing the numbers of newly recruited people within different distances and with different degrees.

In practice $y^1(\alpha, \beta, d_\infty)$ and $y^2(\alpha, \beta, d_\infty)$ are obtained by taking the time averages of the simulated distributions. Let $\bar{y}^1(\alpha, \beta, d_\infty)$ and $\bar{y}^2(\alpha, \beta, d_\infty)$ stand for the resulting distributions when normalized into probability distributions. The corresponding empirical probability distributions are denoted by y^1 and y^2. Note that these distributions are treated as vectors. The criterion used in Section 6 for fitting the parameters α, β, and d^∞ of the exponential model is

$$\|\bar{y}^1(\alpha, \beta, d_\infty) - y^1\|^2 + \|\bar{y}^2(\alpha, \beta, d_\infty) - y^2\|^2.$$

The role of d^∞ is important for networks involving multiple connected components. For example, it affects the probability at which a company connected to a clique outside of the main component recruits new board members from the companies in the main component. In a sense, it is a virtual distance between two firms that are not connected.

5.2. Leaving vacancies empty

In practice the number of positions in companies varies. There can be several reasons for this. First, when someone leaves a position it may take some time to find a new person. Consequently, the vacancy is temporarily left empty. The second reason is that the number of positions can be increased if an interesting candidate is available. One phenomenon that can be seen in real world data is "board swapping"; a person leaves one position and is immediately given another somewhere else. For example, in the empirical application presented in next section, 4%–8% of yearly board appointments involve persons switching from one board to another.

One way to treat empty positions is simply to assume that no vacancy is left empty. The number of position in a board is taken as the maximum number of board members observed in a period of time. When simulating the model it can be assumed that all these positions are filled as if there was somebody holding each empty seat. This means that when counting the empirical distributions for degrees, an empty position is treated as if it was held by somebody who has no other positions.

6. Empirical Application

6.1. Data

The initial data consists of 826 Finnish companies, their board members and CEOs in each year from 2005 to 2015 as documented in the Finnish Trade Register. The set of companies includes the five hundred largest companies in 2013, 2008, 2003, 1998, and the companies listed in Nasdaq Helsinki in 2013. Majority of companies are non listed; 128 are listed. Due to mergers, acquisitions, and bankruptcies, there are also companies that do not exist throughout the sample period. These companies are simply excluded from the analysis leaving us with 730 firms.

The firms corresponding to the relevant labour market of board professionals is extracted recursively: first all the firms that are in the main component in all years are included, all firms connected to these companies in some year through interlocking board members are included, and so on until no new firms are included. Note that the main component can be different each year and therefore taking all the main components would lead to a slightly different selection. Altogether this selection consists of 484 firms, and it can be considered as the pool of companies acting in the same labour market. In particular, if any of the firms in our sample has recruited a new board member, who is already in a board of some company, it is highly likely that the firm from which the new person is taken is in our sample. Moreover, the selection of the firms that is left outside of our sample is only weakly networked; 86% of these firms are not connected to any other firm, and the largest cliques in this pool contain only three firms. Results for the selection of firms that have been in the main component in each year are presented in Appendix B.

The descriptives statistics of the sample of 484 companies are collected in Table 1. Note that some of the model parameters can be directly obtained from these observations. In particular, the estimated probability P_n that a new board member is an outsider is 80% is obtained by dividing n_e with n_o. The probability of opening a vacancy is on aggregate about 16%, and the largest observed degree in the data is $K = 7$. In our data the number of board members varies between 2 and 33 people median being 12. Recall that the board size is usually set in the corporate by-laws, but if not, then there should be from one to five board members according to the Finnish legislation.

	Value
Number of firms	484
Number of positions	3731
Median board size	12
Median number of empty positions/year	630
Total number of vacancies (n_o)	5787
New board members without a position (n_e)	4639
New board members with a position	1148
Probability of recruiting an outsider (P_n)	80%
Largest observed degree (K)	7
Average clustering coefficient of the firm projection	0.20

Table 1: Descriptive statistics.

What is notable in the descriptive statistics is that only 20% of the new board members are found through the network channel. However, even this fraction is enough to create a rather well connected network containing around 350 companies in the main component each year, while at the same time people in the network change reasonably fast. Note that not all companies in our sample belong to the main component, but each of them is at least "indirectly" linked to the main component when considering the firm links throughout out the period 2005–2015. The majority of new board members, 80% of them, are people who did not have any affiliation in previous year's network. Majority of these people, 98% of them, did not have any affiliation in any prior year when getting nominated.

6.2. Estimated model

The solid line with circles in panel (a) of Figure 3 illustrates the distance degree of newly recruited board members when they are taken from the population of all holders of the available board positions in the network. The dotted line with triangle markers represents the distribution for a model in which the new persons are selected at random, i.e., $\alpha = \beta = 0$ in equations (2) and (3). As can be seen, the observed distribution is considerably skewed towards small distances compared to the distribution in the random model. This indicates that there is a considerable distance effect in filling the vacancies.

The skewness of the distance distribution towards small distances means that it is not only that vacancies are filled from immediate neighborhoods in the network but also from further away in the network. We emphasize that in the literature on social networks it is typical to test the likelihood of a triadic closure, which in our case would mean that a contact of a board member becomes linked to his or her board members, i.e., the new board members comes from distance one. The distance distribution shows that the other distances matter, too, hence giving evidence that the indirect connections are highly relevant in the search of new board members.

The degree distribution of new board members is presented in panel (b) of Figure 3. This distribution has a slightly fatter tail than the corresponding distribution of the random model, indicating that there is a form of preferential attachment in filling the vacancies. The distributions of Figure 3 are the ones that are used in fitting the parameters α, β, and d^{∞} of the exponential search model (2) and (3). The empirical frequencies at which vacancies are opened for positions with different degrees are used as probabilities $P_o(n_v)$; these are $(0.16, 0.14, 0.15, 0.13, 0.11, 0.13, 0.19)$ for degrees one, two, and so on. Apart from degrees 3, 6, and 7, these differ statistically significantly from the empirical frequency for the whole population 15.5%.[‡] The rest of the estimated parameters are collected in Table 6.2. Standard errors were obtained by parametric bootstrapping; 200 samples of data were simulated, and the model was refitted to these networks.

When comparing the empirical and simulated distributions, it can be seen that the fitted model performs well. The simulated distributions of distances and degrees of new board members are the gray curves in Figure 3. The non-monotonicity of the distance distribution is an interesting detail which

[‡]For degrees 6 and 7 we observe less than 20 vacancies, which explains why the probabilities 0.13 and 0.19 are not statistically different from 0.155 in a one-way Chi-squared test.

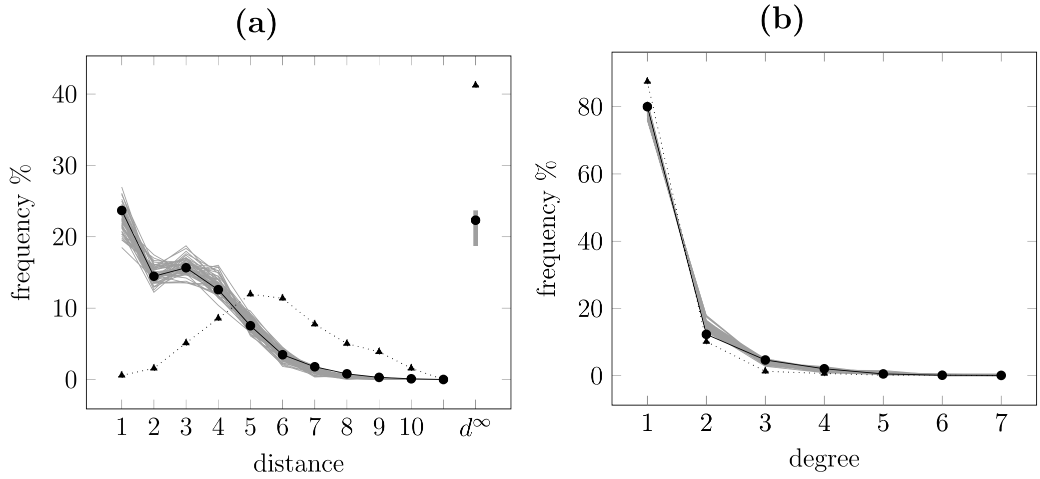


Figure 3: (a) Distance distribution of new board members (solid/circle) and 50 simulated distributions corresponding to the fitted model (gray) and a distribution with randomly filled vacancies (dotted/triangle). (b) Degree distribution of new board members and 50 simulated distributions corresponding to the fitted model (gray) and a model with randomly filled vacancies (dotted).

	Value	SE	95% CI	Interpretation
α	0.79	0.06	$(0.69, 0.90)$	Distance parameter in equation (2)
β	1.29	0.24	$(0.95, 1.50)$	Degree parameter in equation (3)
d^∞	11.5	0.89	$(9.9, 11.9)$	Distance between unconnected positions

Table 2: Parameter estimates, standard errors (SE), and 95% confidence intervals (95% CI).

results from the network being composed of several components that are not connected to each other in each period. In particular, firms outside of the main component are typically members of small cliques of companies, and therefore when searching for new board members among their neighbours, they tend to find them very close in the network. The parameter d^{∞} plays an important role for these companies, because it determines the probability of recruiting people outside of the boards in their neighbourhoods. In particular, d^{∞} affects directly to the probability of person being recruited outside of the clique to which the company belongs to. Second, this parameter affects the non-monotonicity of the distance distribution. The higher the value of d^{∞} the less likely it is that firms within small cliques appoint persons found in the main component, and rather appoint persons found in the same clique, i.e., within distance one. Hence, increasing d^{∞}, increases the peak at distance one.

In addition to matching well to the empirical distributions over degrees and distances of new board members, the other characteristics of the fitted model are reasonably close to empirical observations. For example, when comparing the simulated degree distribution of the fitted model and the empirical distribution in terms of frequencies, they are close to each other (the square error is 0.001 for the probability distributions of degrees). The average clustering coefficient of the firm projection of the fitted model is about 0.17, while the average over the observed networks is 0.20.

The main lesson from our empirical application is that the indirect connections matter a lot when new board members are searched through the network channel, and the new board members are not necessarily chosen from nearby neighborhoods in the network. This supports the hypotheses that the network transmits information regarding potential board members rather than serves the purposes of favor exchange. Comparisons to the specification where only short distances would matter in the selection are presented in Appendix B. The shape of the distance distribution indicates that signals on the quality of people diffuse at longer distances in the network. Although our application is rather specific, it is likely that this distance effect is a common feature of the network search channel in other labor markets as well. Moreover, we believe that similar patters are present in the corporate networks of other countries, too. Many of the firms in our sample are internationally networked both through ownership and board members. For example, our rough estimate for the number of foreign board members in our data is 20%. Hence, our findings are likely to reflect international recruitment practices.

7. Conclusions

In this work we have formulated a tractable model for analyzing the network search channel in recruiting new people. In this paper the search model is presented for two-mode networks where people are connected to each other by sharing the same affiliation. This restriction is motivated by our application related to board interlock networks, but we emphasize that the formulation of the model to a more standard one-mode setup is possible. What is important in two-mode networks is that an individual may hold several positions at the same time, and it is exactly the people having multiple positions who create the ties between people in the network. Second, the two-mode structure allows for an easy way to keep track of peoples' affiliations and updating them simply by rewiring a single edge each time when someone is given a new position.

The search model has simple microeconomic foundations: the distance and degree of individuals in the network affect the productivity signal that the companies have on them. The further away a person is from a company, the less reliable the signal, and the higher the degree, the larger the expected productivity. These features can be interpreted in terms of peer referral and preferential attachment, respectively. It is also demonstrated that the model fits well to empirical data and is capable of explaining the formation of board interlock networks. Moreover, the empirical application shows that social networks play an important role when firms hire new board members, and the structure of the board interlock network is shaped by the search process.

Our empirical findings support the hypothesis that board interlocks help in screening, and potentially enhance the firm-board member match. This is in line with the findings indicating that well connected firms tend to perform better than less connected (Larcker et al., 2013). However, the relationship between firm performance, wages, an corporate governance is to be verified for our sample of firms. In particular, if board interlocks exchange favors, we would expect indicators related to the quality of corporate governance to be positively correlated with the value of the firm, and negatively with executive wages. If on the contrary, board interlocks help in screening, as suggested by our findings, we should not expect wages of board members to change with measures of firm governance.

There are several possible extensions of the model, such as allowing for longer memory in the affiliations or firm heterogeneity in their practices of opening and filling vacancies. In particular, there could be a tendency to recruit new board members from same type of companies, see Currarini et al. (2009) on the role of same-type bias in a matching process of friendship network formation. It is possible to include individual and firm level control variables in the exponential search model introduced in this work. In particular, the questions related to the impact of age, experience, and gender on the likelihood of becoming selected to a vacancy are of interest and possible future extensions of the model. Another line of generalizations would be to let the number of positions and firms vary.

Appendix A. Mathematical Proofs

Proof of Proposition 1. To prove the irreducibility it is sufficient to show that it is possible to get from one neighbourhood configuration to any other by opening and filling the vacancies. Hence, take two neighbourhood configurations N^1 and N^2. First, by A1 and A2 there is positive probability that all the positions of N^1 are opened in the first period one and filled with the new individuals. Moreover, there is a positive probability that after filling the vacancies with new individuals, exactly the ones which have holders with multiple positions in N^2 are opened (A1) in the second period. For each vacancy it is possible that any of the holder of other positions becomes the new holder by (A3). Hence, there is a positive probability that the new holders have exactly the same set of positions that the holders of the positions in N^2. □

Proof of Proposition 2. The productivity is $z^j = s^j + \varepsilon^j$, $j = 1, 2$, for signals $s^1 = s\left(n_{v^1}, d_{v^1,v}\right)$ and $s^2 = s\left(n_{v^2}, d_{v^2,v}\right)$. The terms ε^j, $j = 1, 2$, are mean zero error terms. The ex-ante distributions of s^1 and s^2 are the same, while ε^1 is a mean preserving spread of ε^2 by the assumption of the proposition.

Consider the functions $f^j(s) = \mathbb{E}_{\varepsilon^j}[u(z)|s]$, $j = 1, 2$, i.e., the expected utilities for a given signal s and mean zero error terms ε^1 and ε^2. Because the firm is risk averse, it holds that $f^2(s) \geq f^1(s)$. The probability that the holder of v^j is better than the holder of another position is the probability that $f^j(s)$ is higher than $\bar{u}$ which corresponds to the expected value of the other position. Because the ex-ante distributions of signals are the same and $f^2(s) \geq f^1(s)$, it follows that the probability of $f^2(s) \geq \bar{u}$ is no smaller than the probability of $f^1(s) \geq \bar{u}$. Hence, the result follows. □

Proof of Proposition 3. To suppress the notation let us denote the signals by $s^1 = s\left(n_{v^1}, d_{v^1,v}\right)$ and $s^2 = s\left(n_{v^2}, d_{v^2,v}\right)$. Recall that once the signal is realized the productivity is $z^j = s^j + \varepsilon^i$, $j = 1, 2$. Given that the ex-ante distributions of signals are the same, the first order stochastic dominance implies that ε^2 has the first order stochastic dominance over ε^1. By definition this means that for any increasing u it holds that $f^2(s) \geq f^1(s)$ for all s, where $f^j(s) = \mathbb{E}_{\varepsilon^i}[u(z)|s]$, $j = 1, 2$. As in the proof of Proposition 2 it follows that the probability of choosing the holder of v^1 is no larger than the corresponding probability for the holder of v^2. □

Proof of Proposition 4. Taking the expected value of the CARA utility over the normal distribution with mean $\mu = A + b(n_i) + \varepsilon_{i,k}$ and variance $\sigma^2 = -2a(d_{i,k})$ yields

$$1 - exp\left(-\gamma\mu + \gamma^2\sigma^2/2\right) = 1 - exp\left[\gamma(-A - b(n_i) - \varepsilon_{i,k}) - \gamma^2 a(d_{i,k})\right].$$

Maximizing this function is equivalent to maximizing

$$\gamma a(d_{i,k}) + b(n_i) + \varepsilon_{i,k}.$$

Because $\varepsilon_{i,k}$, $k \in F$ and $i \in I$, are type I extreme value distributed, the ex ante probability of choosing a candidate is determined by the logit expression (6) for the conditional choice probabilities, see, e.g., McFadden (1981). □

Appendix B. Auxiliary Estimations

In this appendix we estimate the model for alternative sample selections and consider the outcomes when the search in the network is restricted in distance.

First, let us next consider how the results change when the model is estimated for a different selection of firms. The specification analyzed in the main body is referred to as the baseline selection. We compare the baseline selection to the one including firms that have been in the main component of the firm projection throughout the time period 2005–2015 (denoted by $\cap PC$ in Table B.3). To make the parameter estimates of β comparable, K is set to 7 in the estimation. We also tested how the selection of firms would change if we took the union of firms in the main component of each period. This would lead to a selection of 466 firms, i.e., only 14 firms from the baseline selection would be dropped, and the estimations would lead practically to same results as in the baseline case.

Selection	α	β	d^∞	P_n	firms	positions
Baseline	0.79	1.29	11.5	80%	480	3731
∩PC	0.97	1.94	18.8	77%	193	1783

Table B.3: Parameter estimates for different selections of firms.

As can be observed from the estimated parameters, both the distance and degree terms are higher in the ∩PC case compared to the baseline selection. This means that the firms who are better networked put also stronger emphasis on candidates positions in the network. This is also visible in the distance distribution of Figure B.4. Moreover, these firms also tend to recruit new board members among each others' boards, i.e., they show a form of homophily. Note that P_n, the probability of recruiting people outside of the boards of 193, is 77%, which is smaller than in the baseline selection even though the pool of companies is much smaller.

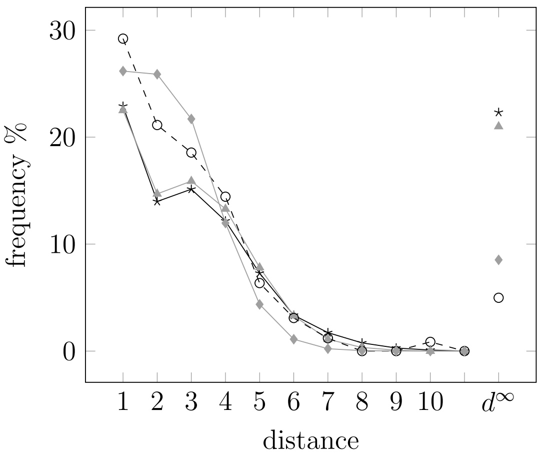


Figure B.4: Distance distributions of new board members in the baseline case; empirical (solid/star) and fitted (gray/triangle), and ∩PC case; empirical (dashed/circle) and fitted (gray/diamond). The fitted distributions are from 50 simulations.

Recall that if the board nominations reflect exchanges of favor, we would expect the short distances to be most important in finding new board members. To test what would happen if the search is restricted firms' neighborhoods within different distances, we fitted the model with the restriction that all people in the network that are further away than $\bar{d}$ are treated as if they could not be reached through the network channel. Figure B.5 shows the distance distributions for the fitted models for different values of $\bar{d}$. As can be seen, for small values of $\bar{d}$ the distance distribution is considerably far from the observed distribution. This supports the hypothesis that the information on possible board members flows through the network.

The reason why there is no peak at distance one when $\bar{d} = 1$ is that if the parameter α was increased, which would mean that people are more likely to come from distance one, then the network would soon fragment such that the peak at distance d_∞ would increase. Hence, a fragmented network with small cliques of firms would possibly indicate that board nominations reflect favor exchanges.

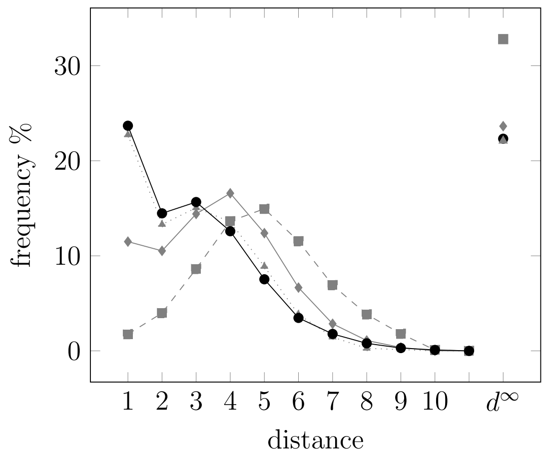


Figure B.5: Distance distributions of new board members in the baseline case; empirical (star/dashed), fitted (triangle/gray), and ∩PC case; empirical (circle/dotted) and fitted (diamond/gray). The fitted distributions are from 50 simulations.

References

A.L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.

J. Berardi, P. Seabright, Professional networks and their coevolution with executives careers: Evidence from europe and the US, Working paper (2011).

S.A. Boorman, A combinatorial optimization model for transmission of job information through contact networks, Bell Journal of Economics 6 (1975) 216–249.

Y. Bramoullé, S. Currarini, M. Jackson, P. Pin, B.W. Rogers, Homophily and long-run integration in social networks, Journal of Economic Theory 147 (2012) 1754–1786.

S. Brenner, The risk preferences of U.S. executives, Management Science 61 (2015) 1344–1361.

A. Calvó-Armengol, Job contact networks, Journal of Economic Theory 115 (2004) 191–206.

A. Calvó-Armengol, Y. Zenou, Job matching, social network and word-of-mouth communication, Journal of Urban Economics 57 (2005) 500–522.

S. Cheng, Board size and the variability of corporate performance, Journal of Financial Economics 87 (2008) 157–176.

S. Currarini, M.O. Jackson, P. Pin, An economic model of friendship: Homophily, minorities, and segregation, Econometrica 77 (2009) 1003–1045.

D. Czarnitzki, K. Kraft, Capital control, debt financing and innovative activity, Journal of Economic Behavior and Organization 71 (2009) 372–383.

P. Diamond, Mobility costs, frictional unemployment, and efficiency, Journal of Political Economy 89 (1981) 798–812.

C. Dustmann, A. Glitz, U. Schönberg, H. Brücker, Referral-based job search networks, Review of Economic Studies 83 (2016) 514–546.

J. Engelberg, P. Gao, C.A. Parsons, The price of a CEO's rolodex, The Review of Financial Studies 26 (2013) 79–114.

T.S. Evans, A.D.K. Plato, Exact solution for the time evolution of network rewiring models, Physical Review E 75 (2007) 056101.

E. Fich, A. Shivdasani, Are busy boards effective monitors?, The Journal of Finance 61 (2006) 689–724.

F. Fontaine, Why are similar workers paid differently, Journal of Economic Dynamics and Control 32 (2008) 3960–3977.

J. Galaskiewicz, S. Wasserman, A dynamic study of change in a regional corporate network, American Sociological Review 46 (1981) 475–484.

M. Galenianos, Learning about match quality and the use of referrals, Review of Economic Dynamics 16 (2013) 668–690.

A. Galeotti, L.P. Merlino, Endogenous job contact networks, International Economic Review 55 (2014) 1201–1226.

M.S. Granovetter, The strength of weak ties, American Journal of Sociology 78 (1973) 1360–1380.

M.S. Granovetter, Getting a Job: A Study of Contacts and Careers, Harvard University Press, Cambridge, 1974.

K. Hallock, Reciprocally interlocking boards of directors and executive compensation, Journal of Financial and Quantitative Analysis 32 (1997) 331–344.

Y.M. Ionnides, D.L. Loury, Job information networks, neighborhood effects, and inequality, Journal of Economic Literature 42 (2004) 1056–1093.

M.O. Jackson, A survey of network formation models: Stability and efficiency, in: G. Demange, M. Wooders (Eds.), Group Formation in Economics: Networks, Clubs, and Coalitions, Cambridge University Press, Cambridge, UK, 2005.

M.O. Jackson, B.W. Rogers, Meeting strangers and friends of friends: How random are social networks?, The American Economic Review 97 (2007) 890–915.

J. Koskinen, C. Edling, Modelling the evolution of a bipartite network — peer referral in interlocking directorates, Social Networks 34 (2012) 309–322.

F. Kramarz, D. Thesmar, Social networks in the boardroom, Journal of the European Economic Association 11 (2013) 780–807.

F. Lafond, Self-organization of knowledge economies, Journal of Economic Dynamics and Control 52 (2015) 150–165.

L.M. Lalanne, P. Seabright, The old boy network: Gender differences in the impact of social networks on remuneration in top executive jobs, Working paper (2016).

D.F. Larcker, E.S. So, C.C.Y. Wang, Boardroom centrality and firm performance, Journal of Accounting and Economics 55 (2013) 225–250.

Y. Liu, Outside options and CEO turnover: The network effect, Journal of Corporate Finance 28 (2014) 201–217.

D. McFadden, Econometric models of probabilistic choice, in: C. Manski, D. McFadden (Eds.), Structural Analysis of Discrete Data with Econometric Applications, MIT Press, Cambridge, MA, 1981.

B. Mintz, M. Schwartz, Interlocking directorates and interest group formation, American Sociological Review 46 (1981) 851–869.

M.S. Mizruchi, What do interlocks do? An analysis, critique, and assessment of research on interlocking directorates, Annual Review of Sociology 22 (1996) 271–298.

J.D. Montgomery, Social networks and labor-market outcomes: Toward an economic analysis, The American Economic Review 81 (1991) 1408–1418.

J.D. Montgomery, Job search and network composition: Implications of the strength-of-weak-ties hypothesis, American Sociological Review 57 (1992) 586–596.

D.T. Mortensen, T. Vishwanath, Personal contacts and earnings: it is who you know!, Labour Economics 1 (1994) 187–201.

M.E.J. Newman, The structure and function of complex networks, SIAM Review 45 (2003) 167–256.

F.H. Page, Jr, M. Wooders, Networks and clubs, Journal of Economic Behavior & Organization 64 (2007) 406–425.

F.H. Page, Jr, M. Wooders, Club networks with multiple memberships and noncooperative stability, Games and Economics Behavior 70 (2010) 12–20.

M.A. Porter, P.J. Mucha, M.E.J. Newman, C.M. Warmbrand, A network analysis of committees in the U.S. House of Representatives, Proceedings of the National Academy of Sciences 102 (2005) 7057–7062.

A. Rees, Information networks in the labor market, The American Economic Review 56 (1966) 559–566.

G. Saloner, Old boys networks as screening mechanisms, Journal of Labor Economics 3 (1985) 255–267.

T.A.B. Snijders, The statistical evaluation of social network dynamics, Sociological Methodology 31 (2001) 361–395.

G. Topa, Labor markets and referrals, in: J. Benhabib, A. Bisin, M.O. Jackson (Eds.), Handbook of Social Economics, Elsevier Science, Amsterdam, 2011.

B. Uzzi, J. Spiro, Collaboration and creativity: The small world problem, American Journal of Sociology 111 (2005) 447–504.

P. Wang, K. Sharpe, G.L. Robins, P.E. Pattison, Exponential random graph (p^*) models for affiliation networks, Social Networks 31 (2009) 12–25.

S. Wasserman, P. Pattison, Logit models and logistic regression for social networks: I. an introduction to markov graphs and p^*, Psychometrika 61 (1996).

The **Aboa Centre for Economics (ACE)** is a joint initiative of the economics departments of the Turku School of Economics at the University of Turku and the School of Business and Economics at Åbo Akademi University. ACE was founded in 1998. The aim of the Centre is to coordinate research and education related to economics.

Contact information: Aboa Centre for Economics, Department of Economics, Rehtorinpellonkatu 3, FI-20500 Turku, Finland.

www.ace-economics.fi

ISSN 1796-3133